DOCUMENT RESUME

ED 108 661 IR 002 186

TITLE	Installation Manual for the Uniterm System of Coordinate Indexing.
INSTITUTION	Documentation Inc., Washington, D.C.
SPONS AGENCY	Department of Defense, Washington, D.C.
REPORT NO	PB-126-671
PUB DATE	Oct 53
NOTE	31p.
AVAILABLE FROM	Photoduplication Service, Library of Congress, 10 First Street, S.E., Washington, D.C. 20540 (PB 126 671, Photocopy $8.00; Microfiche $4.00)
EDRS PRICE	MF-$0.76 PLUS POSTAGE. HC Not Available from EDRS.
DESCRIPTORS	*Coordinate Indexes; *Indexing; *Information Retrieval; *Information Storage; Manuals; Search Strategies; Subject Index Terms
IDENTIFIERS	*Uniterm System

ABSTRACT

The Uniterm system of coordinate indexing, developed by the Armed Services Technical Information Agency, provides a method for the organization, storage, and retrieval of information. The contents of reports are analyzed, and terms are assigned to the documents. For each term there is a card upon which the accession number of documents assigned that term are entered. Reports may be retrieved by looking up the appropriate term or by comparing the document numbers on two or more term cards. Thus, the identical numbers appearing on the three cards "air" and "ducts" and "icing" would lead to documents on air duct icing. The document may then be retrieved directly, or a card containing bibliographic information, abstract, and uniterms may be consulted. Rules and directions for indexing and posting are given, along with some references and search methods. (LS)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

CB126-671 TR

ED108661

ARMED SERVICES TECHNICAL INFORMATION AGENCY

INSTALLATION MANUAL FOR THE *Uniterm* SYSTEM OF COORDINATE INDEXING

OCTOBER 1953

PUBLISHED BY DOCUMENT SERVICE CENTER
KNOTT BUILDING, DAYTON 2, OHIO

IR 002 186

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

THE ARMED SERVICES TECHNICAL INFORMATION AGENCY

INSTALLATION MANUAL

for the

UNITERM SYSTEM OF COORDINATE INDEXING

Prepared by Documentation Incorporated
under Contract AF 18(600)-376

OCTOBER 1953

Published by:

DOCUMENT SERVICE CENTER
Knott Building
Dayton 2, Ohio

Introduction

The Uniterm System of Coordinate Indexing is a scientifically developed, simple and efficient method for the organization, storage and retrieval of information. Essentially it consists of analyzing the contents of reports into a basic vocabulary together with a method of retrieving their contents by any element or combination of elements in the vocabulary. As a result, it provides a degree of search effectiveness unobtainable through conventional library catalogs. The System was developed under an intensive research program conducted by Documentation Incorporated for the Armed Services Technical Information Agency (ASTIA).

The features of compactness, flexibility, economy of installation, ease of maintenance and ease of searching make the Uniterm System especially suitable for the control of technical documents whose full subject content is frequently lost because of difficulty in gaining catalog control through conventional means. The advantages of the System are derived from the simple principles which are introduced in the following pages.

CONTENTS

SECTION		
I	General Description	1
II	Uniterm Indexing with ASTIA Cards	11
III	Indexing and Posting	14
IV	Rules of Coordinate Indexing	19
V	Reference Work with a Coordinate Index	24

Additional copies of this manual may be obtained from ASTIA Document Service Center, Knott Building, Dayton 2, Ohio.

SECTION I

GENERAL DESCRIPTION

Language is based on words. Many ideas and concepts are expressed in language as combination of words. The number of such combinations for any body of information may be large, but the number of individual vocabulary terms used is often surprisingly small. For example, the number of chemical elements is small compared to the number of compounds.

In the Uniterm System of Coordinate Indexing, the information contained in a collection of documents is analyzed into the simplest practical word units of information -- hence, "Uniterm". Each Uniterm is assigned a separate card. Each document is assigned a serial number, and the number is posted on all cards headed by the Uniterms by which the document has been analyzed. The word units usually consist of single-word ideas or concepts, each of which is associated with a body of related information.

For example, the word "air" is associated with a number of different ideas in the phrases, "air ducts", "air speed", "cold air", etc. There are two ideas or terms in the phrase, "air ducts" -- one the class of <u>air</u> ideas, and the other the class of <u>duct</u> ideas. When we use the two terms together, we perform the operation known as logical conjunction. This means simply the combination of "air" <u>and</u> "ducts",

and it is obvious that the same concept can be expressed as "ducts" and "air". Thus the order of terms is immaterial since the same total body of information is represented by either arrangement as in the following diagram:

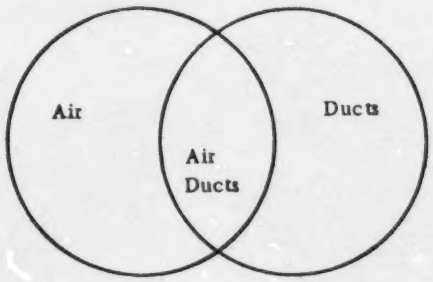

Figure 1.

The overlapping or common area of "air" and "ducts" represents a smaller body of information than either of the individual areas. It is easy to see that a more specific area can be obtained by using a larger number of terms in conjunction, thus narrowing the field of search as in the following diagram:

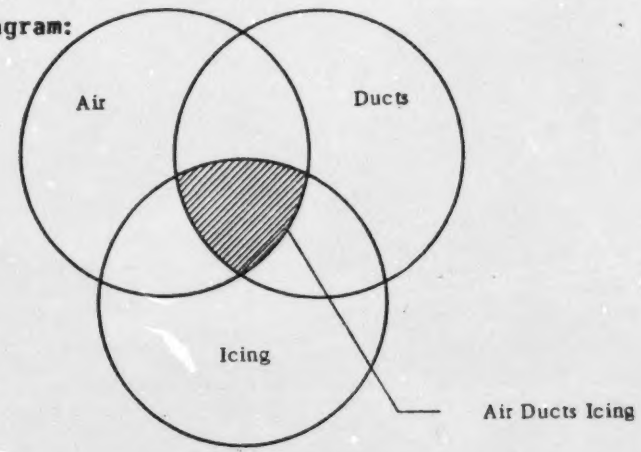

Figure 2.

Since the same logical conjunction is obtained regardless of the order of the terms, the three separate unit terms, "air", "ducts", and "icing", when coordinated, are equivalent to the six possible ways by which this information can be represented in conventional indexing systems:

AIR DUCTS - ICING

DUCTS, AIR - ICING

ICING - AIR DUCTS

ICING - DUCTS, AIR

AIR - ICING - DUCTS

DUCTS - ICING - AIR

Simplicity of Catalog Form

In the manual Uniterm System, as explained in detail in the following pages, this logical conjunction is achieved by searching for the document numbers which are common to two or more Uniterm cards. There is no reason to assume that "air" is any more important than "ducts" or "icing" and consequently no reason to subordinate any one term with respect to any other term as is done in conventional systems of indexing and classification. The making of such assumptions

at different times for different cases gives rise to elaborate cross reference structures of such steadily increasing complexity that the conventional systems eventually either become meaningless or make extreme demands upon personnel staffs in maintaining the systems' integrity. By its planned indifference to the order of words, the Uniterm System eliminates the need for conventional cross references.

Economy of Catalog Space

Because of the elimination of permutations and combinations of basic ideas used in most present-day systems, a dramatic saving in catalog space results from the use of the Uniterm System. The size of the System is a function of the number of vocabulary terms rather than of the number of expressions of ideas generated by the combination of terms. For a collection the size of ASTIA's, this reduction of catalog size is in the order of 90%.

After a body of information has been analyzed into the basic ideas or Uniterms, the number of terms increases very slowly as the system grows. However, the number of permutations and combinations of these new basic terms increases rapidly and exponentially, in fact by (2^n-1), where n is the number of Uniterms.

Operational Benefits

The indexer is the first to benefit from the elimination of the permutations and combinations of words found in conventional systems. Freed from the time-consuming necessity of making subjective decisions concerning the order of words and of creating complicated and highly specific indexing phrases and cross references, he assigns Uniterms rapidly and liberally, laying the foundation for a depth of analysis hitherto impractical. The poster, who transfers numbers to Uniterm cards, benefits because his task is simple compared to the corresponding work of preparing and filing full catalog cards for each assignment in a conventional system. The searcher benefits because he can address the coordinate index from any point of view because he can be as specific or as general as he wishes and because he need not wonder about the form of expression used by the indexer.

Manual and Machine Manipulation

The logical fragmentation of catalog elements inherent in the Uniterm System makes retrieval of information possible by machine or manual methods. Currently under development is a machine which will permit random filing and

retrieval of Uniterm cards for posting and searching. Also being developed is a completely automatic indexing machine which will present the document numbers in answer to reference questions typed out on a conventional typewriter keyboard.

While the magnitude of the overall problem of information control made the need for machine systems apparent many years ago, the search for a manual method was renewed with the realization that there are many small installations in need of an improved manual method of information control.

Physical Appearance

The Uniterm System in its manual form uses cards of any convenient size, such as 3" x 5", 4" x 6", or 5" x 8", in conventional catalog card trays, visible index files, or looseleaf volumes. The cards have spaces at top and bottom for the Uniterms and are divided vertically into ten columns, numbered from zero through nine, in which document accession numbers are posted. The columnar arrangement is designed simply to break up the mass of numbers posted on the cards. The Uniterm card* looks like this in use:

*Uniterm cards can be printed locally by the user, or, as a matter of information, they can be obtained from Documentation Incorporated, 2521 Connecticut Avenue, N. W., Washington 8, D. C. at a cost of $15.00 per thousand for the recommended 5" x 8" size.

TERM Ducts

0	1	2	3	4	5	6	7	8	9
	11	32	103	24	75	206	37	98	
	61	122	113	34	355	216	47	208	
		312	643		755	826	447		
					785				

Note that the last digit of each serial accession number corresponds to the number of the column in which it is posted. Note also that since documents are cataloged in order of their serial accession numbers, no space need be left in any column, since new documents will always bear a higher number than any number previously posted.

UNITERM card
TERM

© DOCUMENTATION Incorporated

Figure 3.

Ease of Search

To demonstrate the simple practice of logical conjunction of Uniterms to find reference to wanted information, the following sample cards correspond to the "air", "duct", "icing" example described in Figure 1 and 2 on page 2.

TERM Air

0	1	2	3	4	5	6	7	8	9
110	11	362	13	14	55	186	77		39
230	421	372	73	34	135	216	327		189
		882	133	454	735	876	447		199
				634	785		677		219

TERM Ducts

0	1	2	3	4	5	6	7	8	9
230	11	32	103	24	75	206	37	98	
	61	122	113	34	355	216	47	208	
		312	643		755	826	447		
		882			785				

TERM Icing

0	1	2	3	4	5	6	7	8	9
230	61	112	93	14	35	216	87		
410	421	312	133	434	425				
	721	372	143	554	785				
		882		674					

UNITERM card
TERM

© DOCUMENTATION Incorporated

Figure 4.

It is indicated here that in our body of literature there are twenty-four documents containing information on "ducts"; twenty-nine on "air"; and twenty-one on "icing". Some of the "duct" information may be on gas ducts, water ducts, and air ducts. We are concerned with finding those having to do with "air ducts icing". The essence of any successful search is to narrow the field of search. Thus, as in Fig. 1, the area of overlap of duct and air information defining that on "air ducts" is determined by the numbers common to both Uniterm cards. The arrangement of numbers by final digits and in ascending order makes it easy to determine that document numbers 230, 11, 882, 34, 785, 216, and 447 pertain to "air ducts". The field of search is already narrowed to seven documents. By comparing these seven document numbers with the numbers on the "icing" Uniterm card, it is determined the four documents (numbers 230, 882, 785, and 216) contain information on "air ducts icing" (see Figure 2).

Basic Factors for Ease of Search

There are several factors which contribute to the ease of retrieving information from the Uniterm System of Coordinate Indexing. The index headings are chosen and recorded as simple "Uniterms"; the documents are identified in the simplest manner - by serial accession number; the numbers are recorded in ascending

order on Uniterm cards. Retrieval of information is likewise accomplished in the simplest manner - by visual determination of numbers common to two or more Uniterm cards and by rapidly narrowing the field of search with each comparison made. It will be noted that the time of search is determined by the smallest group of numbers to be compared.

The foregoing operation, known as logical conjunction, is the reversal of the process of fragmentation which occurred when the Uniterms were assigned. It can be seen readily that reference questions involving search technique by alternation or negation (such as either "air" or "ducts" or both; or all documents on "icing" with the exception of "ducts") can be handled with equal facility.

SECTION II

UNITERM INDEXING WITH ASTIA CARDS

For those users of ASTIA services who may wish to adopt the Uniterm System in controlling ASTIA materials, Uniterm tracings will be included on regular ASTIA catalog cards in addition to conventional headings. This manual gives instructions for the installation and use of the Uniterm System utilizing ASTIA cards as document references. Additional instructions are presented so that non-ASTIA material can be processed to provide an integrated file of holdings.

Beginning with ASTIA document number 15001, ASTIA catalog cards will contain Uniterm tracings on the reverse side of the card. These cards will look like this:

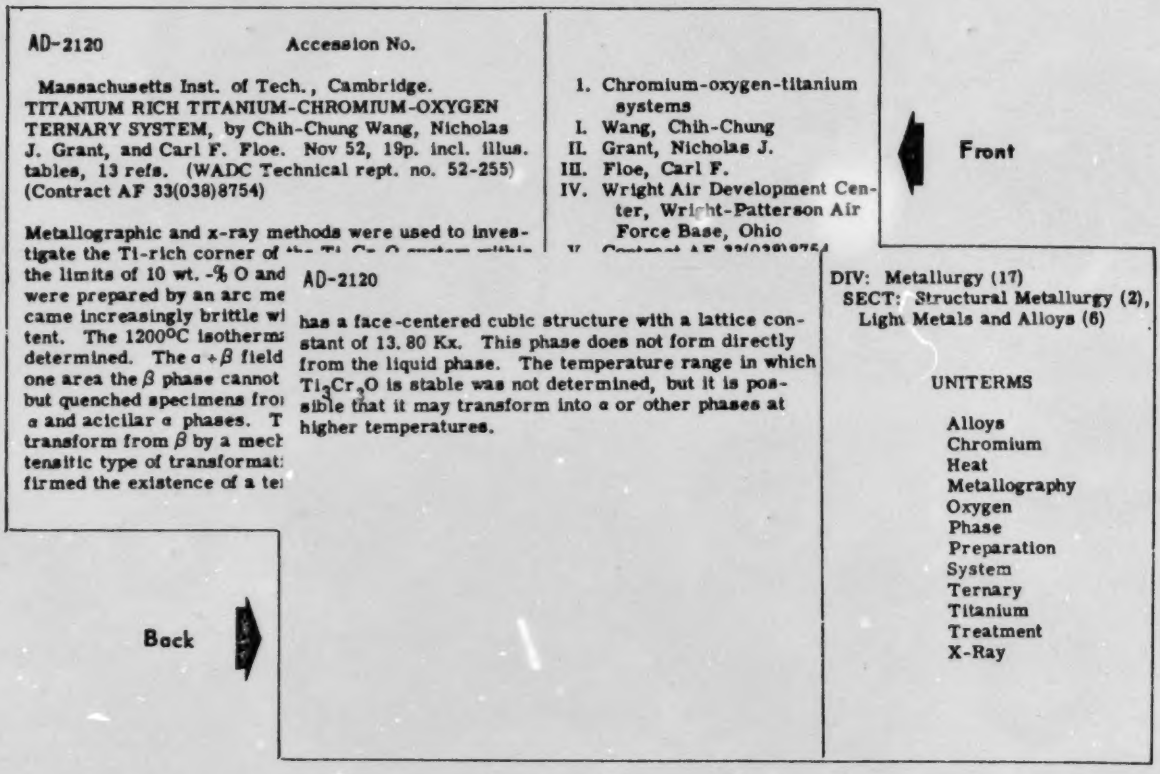
AD-2120
Accession No.
Massachusetts Inst. of Tech., Cambridge.
TITANIUM RICH TITANIUM-CHROMIUM-OXYGEN TERNARY SYSTEM, by Chih-Chung Wang, Nicholas J. Grant, and Carl F. Floe. Nov 52, 19p. incl. illus. tables, 13 refs. (WADC Technical rept. no. 52-255) (Contract AF 33(038)8754)
Metallographic and x-ray methods were used to investigate the Ti-rich corner of
the limits of 10 wt. -% O and
were prepared by an arc me
came increasingly brittle wi
tent. The 1200°C isotherm
determined. The α + β field
one area the β phase cannot
but quenched specimens fro
α and acicular α phases. T
transform from β by a mech
tensitic type of transformat
firmed the existence of a te
1. Chromium-oxygen-titanium systems
I. Wang, Chih-Chung
II. Grant, Nicholas J.
III. Floe, Carl F.
IV. Wright Air Development Center, Wright-Patterson Air Force Base, Ohio
Front
AD-2120
has a face-centered cubic structure with a lattice constant of 13.80 Kx. This phase does not form directly from the liquid phase. The temperature range in which Ti_3Cr_2O is stable was not determined, but it is possible that it may transform into α or other phases at higher temperatures.
DIV: Metallurgy (17)
SECT: Structural Metallurgy (2), Ligh Metals and Alloys (8)
UNITERMS
Alloys
Chromium
Heat
Metallography
Oxygen
Phase
Preparation
System
Ternary
Titanium
Treatment
X-Ray
Back

Figure 5

These cards differ from earlier ASTIA cards in two respects:

1. A space is provided immediately following the AD number for the serial accession number to be assigned by the user.

2. Uniterm tracings are provided in the right margin on the reverse of the cards; they represent ASTIA's Uniterm subject analysis of the document.

Catalog cards produced by the user or by any other activity, such as the Atomic Energy Commission, the National Advisory Committee for Aeronautics, and the Library of Congress, can be integrated with ASTIA cards by assigning each card its own accession number. Any activity starting a coordinate index should assign number 1 to the first card placed in the accessions file, number 2 to the second card, number 3 to the third card, and so on, regardless of the ASTIA document number or document numbers of any other card-producing agency.

It is recommended that all the materials covered in a single integrated Uniterm System be stored in order by the serial accession numbers assigned by each user in which event the accessions file serves also as the shelflist. If the materials must be filed in some other order, such as by source, subject classification, ASTIA number, etc., location symbols must be shown on the cards in the accessions file; the need for

shelflists for other arrangements must be decided without reference to the Uniterm System. Cards can be added to the accessions file without adding the materials to the collection, for subject headings and classification numbers can be converted easily to Uniterms without reference to the materials, especially when titles and abstracts are used to aid in the conversion.

SECTION III

INDEXING AND POSTING

The rules of the Uniterm System of Coordinate Indexing are given in Section IV of this manual. These fifteen rules should be studied carefully by those engaged in both indexing and posting. It is always desirable, of course, for searchers to be familiar with the rules even though they not be regular reference workers.

Indexing

Actually, the indexer is concerned primarily with Rules 1 and 2 which simply tell him to determine the key words representing the subject content of the document and how to record these key words for posting. The indexer is not required to create and maintain a list of approved Uniterms, for such a list is created and maintained as the coordinate index itself. Only rarely does the indexer find it necessary to consult the coordinate index as he assigns Uniterms to the accessioned materials. Consequently, there is a great freedom in amending the vocabulary to describe the changing characteristics of a field as reflected in its literature and to continually enhance the effectiveness of the coordinate index without altering previous entries or the structure of the index. These features, plus the simplicity of the rules, make the indexing operation

much more rapid and efficient with the Uniterm System than with any other method.

Posting

The posting of serial accession numbers on Uniterm cards is equivalent to preparing and filing cards in conventional catalogs. In both systems care must be exercised to achieve accuracy. The posting operation is the tedious part of maintaining a Uniterm System, although experience has demonstrated that with proper organization of the work it is less tedious and more economical than card preparation and filing.

The poster has but one major decision to make: that is, whether a term is "free" or "bound" when it is used in the filing position. Rules 3, 4, and 5 in Section IV govern this determination. Generally speaking, words susceptible to combination with other terms in a system are "free" and those found in combination with only one other word in the system are "bound". Bound terms tend to become free terms as a system grows. Therefore, most ASTIA Uniterms will be free terms.

The poster will also be concerned with word form, i.e., singular or plural, substantive or attributive, synonymous and homonymous, etc. Rules 6 through 15 cover these points.

Neatness and accuracy are very important in posting. Careful hand numbering or typing may be satisfactory. Ring-style rubber stemps and numbering machines with repeat settings minimize transcription errors. Small adding or bookkeeping machines may be most satisfactory for large installations, for they can be secured with repeat settings, horizontal tabulation for the columns, easy vertical alignment for the lines, and carriages to suit the width of card; and they produce legible results rapidly.

The Uniterm System of Coordinate Indexing is adaptable to a great variety of conditions. In the places where coordinate indexes are now in operation, no two exactly resemble each other. Because an index is literally tailored to the information collection of each organization, "free" terms in one vocabulary are not "free" terms in another, and so with phrases and inversions. Some include authors' surnames; others do not.

Growth of Uniterm Vocabulary

It is useful to know what to expect during the early growth of a coordinate index. Experience has shown that the number of new Uniterms grows rapidly at the beginning and can be expected to exceed the number of documents indexed. At some point, perhaps between 1000 and 3000 documents, the number of documents equals the number of Uniterms and soon exceeds the

number of Uniterms, since the vocabulary is stabilizing at a slow rate of growth. Therefore, in the very early phases of Uniterm systems there are fewer cards than documents, whereas the reverse is true in the case of the conventional subject heading or a classified catalog.

Since an average of eight Uniterms may be assigned to each document, the growth in assignments may reach many times the growth in Uniterms. Figure 6 shows the growth of Uniterms and Uniterm assignments for a collection of 3000 documents.

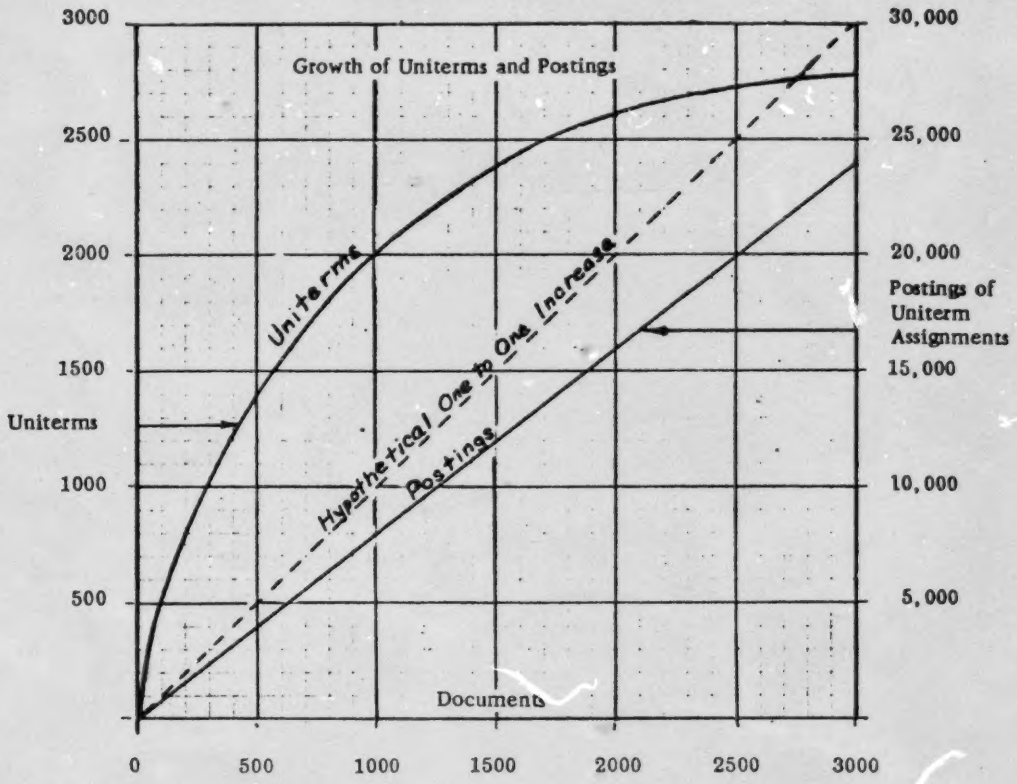

Figure 6.

Any installation can compare its experience with this graph by plotting a cumulative record of its new Uniterms and assignments against the number of documents indexed.

SECTION IV

RULES OF COORDINATE INDEXING

1. **Determine the key words which represent the subject content of the item being indexed.**

   Key words are not limited to single common words in the English dictionary. They include phrases, Arabic numerals, alphabetical characters, model and project numbers, such as AN/ARN-5 and MX-772, in addition to the nouns, adjectives, gerunds, participles, and proper names which make up the bulk of any Uniterm vocabulary.

2. **Record the key words on the serial accessions card so that every retrieval word is a filing word in the posting operation.**

   This is accomplished by putting every word in a phrase in the filing position; thus, if the phrase "gas turbines" is to be used, "turbines" must be recorded in the filing position as well as "gas". The poster then ascertains whether the "turbines" card in this situation bears the single word, "turbines", or must include "gas" and consequently read, "turbines, gas"; also whether the "gas" card bears the single word or the phrase, "gas turbines" (see Rules 3, 4, and 5).

NOTE: A few words are not considered to be retrieval words and need not be recorded on the serial accessions cards in the filing position. Examples are: "Effects" and "Methods".

3. If a word is used alone as an indexing term, it is called a "free" term. Enter it as a single word on a Uniterm card.

In the case of a document on "Helium in Air Turbines for Gasoline Pumps", "helium" is a "free term". By virtue of being used alone once, it remains a "free" term in the future and may not be used later in the vocabulary as part of a phrase.

4. If a word in the filing position is used with only one other word in the vocabulary, it is a "bound" or "not free" term. Enter each word of a two-word phrase on a separate Uniterm card in the filing position, and add the other word if the first word is a "bound" term.

Using the example in Rule 3, when the poster considers the phrase "air turbines", there are only three Uniterms in the vocabulary: "helium", "air", and "turbines". Since both words of the phrases fit this Rule number 4, "air turbines" is typed on one card and "turbines, air" is typed on another.

5. If a word occurs in the system in combination with two or more words, it is a "free" term. Enter it as a single word on a Uniterm card.

Referring to the example in Rules 3 and 4, the above situation occurs when the poster is ready to add "centrifugal", "gasoline" and "pumps" to the vocabulary. "Pumps" is a "free" term because it occurs in combination with "centrifugal" and "gasoline" but the "centrifugal" card is typed as "centrifugal pumps", and the "gasoline" card as "gasoline pump" in accordance with Rule 4. Note that the first word, the word in the filing position, is the one to consider whether "bound" or "free". "Pumps" is a free term in the filing position, but it is added to "centrifugal" and "gasoline" because there is nothing on the subject of centrifugal in the system except centrifugal pumps, and nothing related to gasoline except gasoline pumps. "Free" terms and "bound" terms can be added freely to "bound" terms, but no terms can be added to "free" terms.

NOTE: As the posting is continued, always in serial number order, more terms are added to the vocabulary and actual associations of words increase rapidly in the system. The effect of Rules 3, 4, and 5 is to "free" most of the terms, so that a coordinate index for several thousand items soon has single words for the majority of Uniterms. If, following the example used above, a report on "air ducts" enters the system, the poster must "free" the word "air" according to Rule 5, because it is associated with "turbines" as well as "ducts". The operation is the simple erasure of "turbines" from the "air turbines" card. Note that the serial numbers are not disturbed when a phrase is shortened to a "free" term.

6. Enter substantive forms in the plural only, provided the singular can be inferred conveniently without broad change of meaning.

Example: Grasses

7. Enter singular forms after plurals where the appearance of both terms facilitates indexing and searching.

Example: Supplies; Supply

8. Enter the singular form of foreign words and add the plural.

Example: Fungus; Fungi
Spectrum; Spectra

9. If the singular form is broadly generic or defines a field and the plural is more substantive, enter the singular and add the plural.

Example: Temperature; Temperatures

10. Add attributive or adjectival forms to the substantive words with which they stand in the closest sense relationship.

Example: Geophysics; Geophysical

Sun; Solar

Hydrostats; Hydrostatic

11. Enter gerundive forms and add the participles.

Example: Sintering; Sintered

Printing; Printed

NOTE: The form of the endings of words (singular, plural, nominative, etc.) does not affect the filing order as in conventional systems, since each word appears only once in the filing position.

12. Consider the first word of proper names and foreign phrases to be a "bound" term and enter the full phrase.

Example: Black Sea

Clostridium botulinum *

Clostridium perfringens *

* (No entries are made for the second word here)

13. Enter accepted chemical terms occurring as one word in the form in which they appear.

Example: Amines

Phenylamines

Diphenylamines

Nitrodiphenylamines

14. Where true synonyms occur, enter the well-known form and make a see reference from the other form.

Example: Petrol see Gasoline

Since the see references are consulted rarely, they may be kept conveniently in another file.

15. Enter homonyms as one Uniterm. Where meaning may prove to be unclear in coordination, show meanings parenthetically on separate cards, particularly where many serial numbers are involved.

Example: Pitch (substance)

Pitch (motion)

Pitch (acoustics)

Pitch (angle)

SECTION V

REFERENCE WORK WITH A COORDINATE INDEX

The visual reference use, or searching, of a coordinate index consists of comparing two or more Uniterm cards to determine numbers common to the cards. The common numbers denote the area of logical conjunction of the elements of the question.

It was explained in Section III that because of the size of the ASTIA collection, most Uniterms on ASTIA cards will be "free" terms. However, many ASTIA users will find that ASTIA "free" terms might properly be "bound" terms in their particular systems in view of their specialized interests. If a search shows that all numbers on the "air" card are common to those on the "duct" card, the Uniterms can be changed to "air ducts" and "ducts, air". The change makes it unnecessary to coordinate again under similar circumstances, for either card shows at a glance that "air" and "duct" are bound terms.

For a variety of reasons, reference workers may wish to prepare a record on cards or slips of all coordinations made for searchers. This record serves as:

1. Call slips for documents and for abstracts to be consulted.
2. A file of answers for cases where frequent coordination of the same terms are made, thus making the repetitious coordination of the same terms unnecessary.
3. A source for frequent tabulations to show the volume and kind of reference work which has been performed, thus providing a record not available in most libraries.

TERM:

0	1	2	3	4	5	6	7	8	9